

Agenda

- Methodology
- Results
- Conclusions / Future Work
- Appendix



Intro

Median US Rent is **28% of Income**. NYC Rent is **37 % of Income**

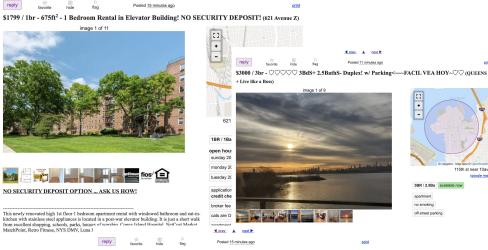
Finding an affordable apartment in NYC can be really hard.

Craigslist is a great place to find affordable apartments



Methodology

- 9,000 Craigslist apt postings (NYC area)
- Used Python, Requests, and Beautiful Soup
- Stored the data with SQLite







CONTACT LANLORD: robertsisman916(@)gmail.com

prev
next ▶

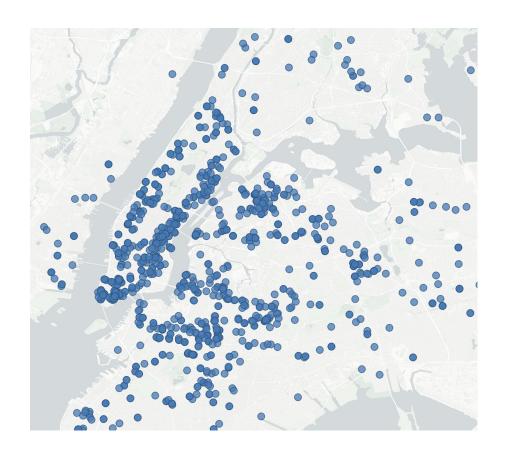
Spacious IBR agartment conveniently located in a well maintained building with a live-in usper. This unit features large living room, king sized bestrown, plenty of natural light through oversized windows, yeared closet space and heauthfully maintained hardwood floors. The building is located only blocks from the IDs & 4 rains and conveniently close to the Metro North (Hudson line) and Major Deegan, as well as some of the best dining, shopping and entertainment in the neighborhood and the convenient in the neighborhood and the convenient of the plant of the convenient of the conven

. do NOT contact me with unsolicited services or offers

Methodology

Cut down data from 9,000 rows to 2,000 rows

Used lat/lon to calculate borough and neighborhood



Methodology

High Variance Data

Best Results Using a Linear Model with Degree 3 Polynomial Transform

Features Price # of Images in Post Latitude / Longitude Sqft # of Bedrooms Dogs Allowed? Cats Allowed? Borough (Dummy Variable)

Results

• Final R² Score: 0.768

 Caveat: High Variance Data Makes Model Unstable

> Not overfit but may be hard to generalize

Need to collect more data

	No Polynomial	Polynomial (degree 3)
Train	0.570	0.806
Validate	0.591	0.817
Test		0.768

Conclusion / Future Work

- Gather More Data
 - Cut down on number of features
 - Use LASSO Cross Validation

- Improve the MVP
 - Site or Chrome Extension for finding good deals on apartments

Thank you.

Appendix

Price and Sqft Distributions

